

Optimizing System Compute and Bandwidth Density for Deployed HPEC Applications

Randy Banton and Richard Jaenicke

Mercury Computer Systems, Inc.

Phone: 978-967-1134

Fax: 978-244-0520

Email: rbanton@mc.com

Email: rjaenicke@mc.com

Topic Area(s):

Embedded Computing for Global Sensors and Information Dominance

Case Study Examples of High Performance Embedded Computing

High-Speed Interconnect Technologies

Abstract: Many high-end deployed military and commercial applications share a common need to achieve high to very high compute and bandwidth density in the smallest possible volume. In addition, deployed military applications layer on additional environmental requirements such as higher levels of shock, vibration, endurance vibration, temperature, and condensing humidity. Each of those adds constraints on the solution space for maximizing compute and communication density.

Not all HPEC applications can use the same solution due to varying limits on total size or weight and varying levels of ruggedness. For example, the size and weight requirements differ greatly for manned surveillance aircraft, large UAVs, and small UAVs. This presentation will explore different options for high-density system designs while meeting the requirements for each of these applications.

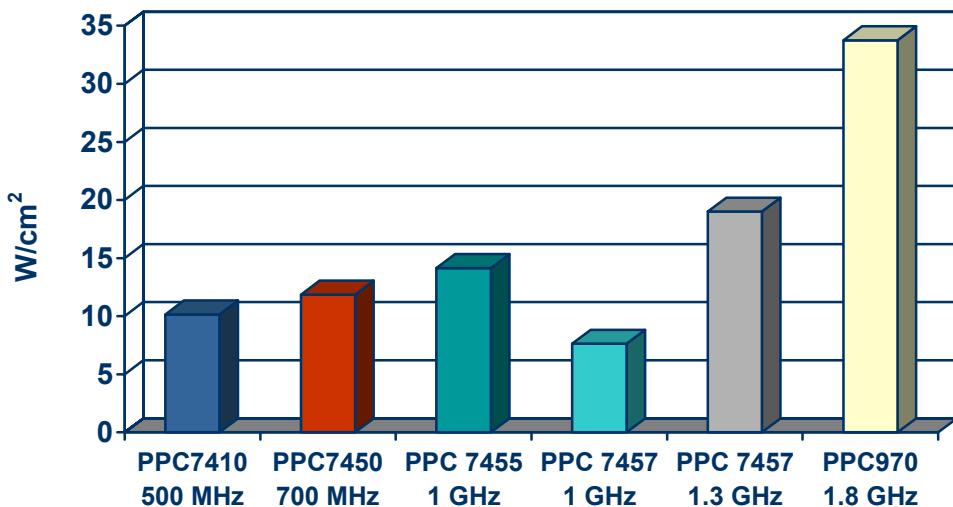


Figure 1. Power flux of recent and near-term PowerPC processors.

Report Documentation Page			Form Approved OMB No. 0704-0188	
<p>Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p>				
1. REPORT DATE 20 AUG 2004	2. REPORT TYPE N/A	3. DATES COVERED -		
4. TITLE AND SUBTITLE Optimizing System Compute and Bandwidth Density for Deployed HPEC Applications			5a. CONTRACT NUMBER	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Mercury Computer Systems, Inc.			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited				
13. SUPPLEMENTARY NOTES See also ADM001694, HPEC-6-Vol 1 ESC-TR-2003-081; High Performance Embedded Computing (HPEC) Workshop(7th). , The original document contains color images.				
14. ABSTRACT				
15. SUBJECT TERMS				
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified	18. NUMBER OF PAGES 19	19a. NAME OF RESPONSIBLE PERSON

A big obstacle is the well-known trend of increasing power consumption for newer faster processors. Faster processors require faster memory, faster interconnect, and more I/O. Not as well known is the power increase due to the move to high-speed switch fabrics clocked up to multiple GHz and memory systems being clocked well into the hundreds of MHz. A high-speed switch fabric and memory system exacerbate the thermal problem even as it solves bandwidth problems. In these cases, it is the large number of components and the concentrated point sources of power that create the biggest challenge beyond that of the overall increasing power and shrinking die size. This combines with the increasing power consumption to create a significant challenge. Figure 1 shows a comparison of past and present processor power flux.

Yet these are not the only demanding components in the system. FPGAs, SRAM, on-board power supply FETs, integrated DC/DC converters, and various ASICs must all be examined as point sources of heat in today's designs. Previous designs have 70% or more of the power in the system going to the processors alone. For current designs, it is typically 40-50%, with the balance going to areas such as interconnect, memory, and supporting components. Large FPGAs used for computing can consume 20W, and the power supply FETs can have the power flux (W/cm^2) of a similar magnitude to the processors.

When component package type is taken into account (e.g., plastic TSOP for DDR-DRAM, PBGA for ASICs), the other components mentioned can approach the same power and thermal management challenge as the processor. Designing a system to optimize the performance density with respect to each of these is challenging.

For any given selection of processor, fabric, and support chips, the areas of system design that can be worked to maximize overall power density, and therefore performance density, are largely driven by the mechanical aspects of the board form factor and the resulting cooling methods for the boards and chassis.

Traditionally, the choices for cooling methods have been convection-cooled (air-cooled) and conduction-cooled. Recent years have seen a growing interest in spray-cooling (evaporation-cooled), because it offers a greater potential for thermal dissipation than either air-cooled or conduction-cooled while providing the benefits of a sealed enclosure for environments with little air, dirt, or corrosive elements. Yet when the size, weight, and maturity of the various cooling methods are examined, air-cooling is shown to meet the right set of tradeoffs against the commercial and military requirements for much of the HPEC space.

This presentation will describe one technique for extracting greater thermal efficiency from an air-cooled design we call "finely managed air." In this approach – covers on the boards, heat-sinks captive in these covers, and airflow shaping in the covers, slots, and chassis inlet areas – are all designed together as a system to carefully direct all available air flow over the hottest components to extract the maximum thermal efficiency. While in conventional board and heat sink design the majority of the air wants to flow around the high-impedance heat sinks instead of through them, the new approach uses features in the cover that directs air through the heat sinks. Since high velocity flow tends to ride up the backplane and starve the front of the board, there is also a need to balance the airflow front to back within a slot to also maximize cooling efficiency.

Features built into the cover achieve this in a way to have the minimum effect on overall pressure drop.

Such “finely-managed air” designs have been verified using both Computational Fluid Dynamics (CFD) modeling and lab measurements of operational hardware at the component, board, and system level, and this paper will present the analysis of that data. In addition to the airflow-shaping aspects of the cover, the cover also adds a very important ruggedizing structural element to the board while not penalizing the available surface area of the PWB with stiffeners, mounting holes, and keep-out regions.

To further maximize the effect of such an approach, each processor or high-power component should get an independent flow of air. Currently high-end deployed commercial off-the-shelf (COTS) high-density designs use large area boards, such as VME 9U x 400mm, to maximize the number of processors per given volume using both the available height and depth. New commercial form factors have similar approaches, such as PICMG3.0 with an 8U high board and less deep at 280mm. In those tall-board designs, components not on the inlet edge of the board will see air that was already heated by processors or other high-power components lower on the board. There can be as many as three to six rows of high-power components in an attempt to maximize processing physically in the layout.

In the face of these thermal considerations, concurrent engineering must be given to board and chassis design. The optimal location for high-power components is along the bottom of the board on the leading edge of the airflow. A long, shorter board therefore provides the maximum thermal dissipation for a system by combining this longer leading edge with a lower flow resistance due to the shorter nature of the board. In order to get the highest density in a given vertical space, two boards stacked vertically with an air intake in the middle provide a significant improvement over past practice in the thermal dissipation capacity for a given board set. An example of such a configuration is shown in Figure 2. For both boards, the hottest components are placed on the leading edge of the airflow.

This organization allows for continued use of a large common backplane. A large common backplane remains the lynchpin to achieving tens of GB/s of inter-chassis bandwidth in a very compact and very high-speed signaling environment.

This presentation will discuss such design trade-offs for different deployed HPEC environments in terms of size, power, and weight constraints.

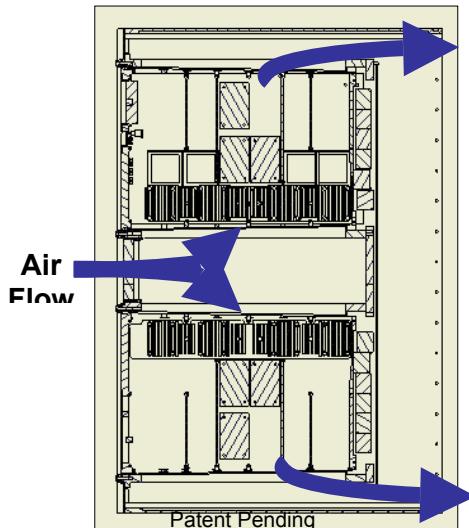


Figure 2. Two short boards mounted vertically sharing the same air intake to minimize vertical space.



Optimizing System Compute Density for Deployed HPEC Applications

Randy Banton, Director, Defense Electronics Engineering
Mercury Computer Systems, Inc.
rbanton@mc.com

Richard Jaenicke, Director, Product Marketing
Mercury Computer Systems, Inc.
rjaenicke@mc.com



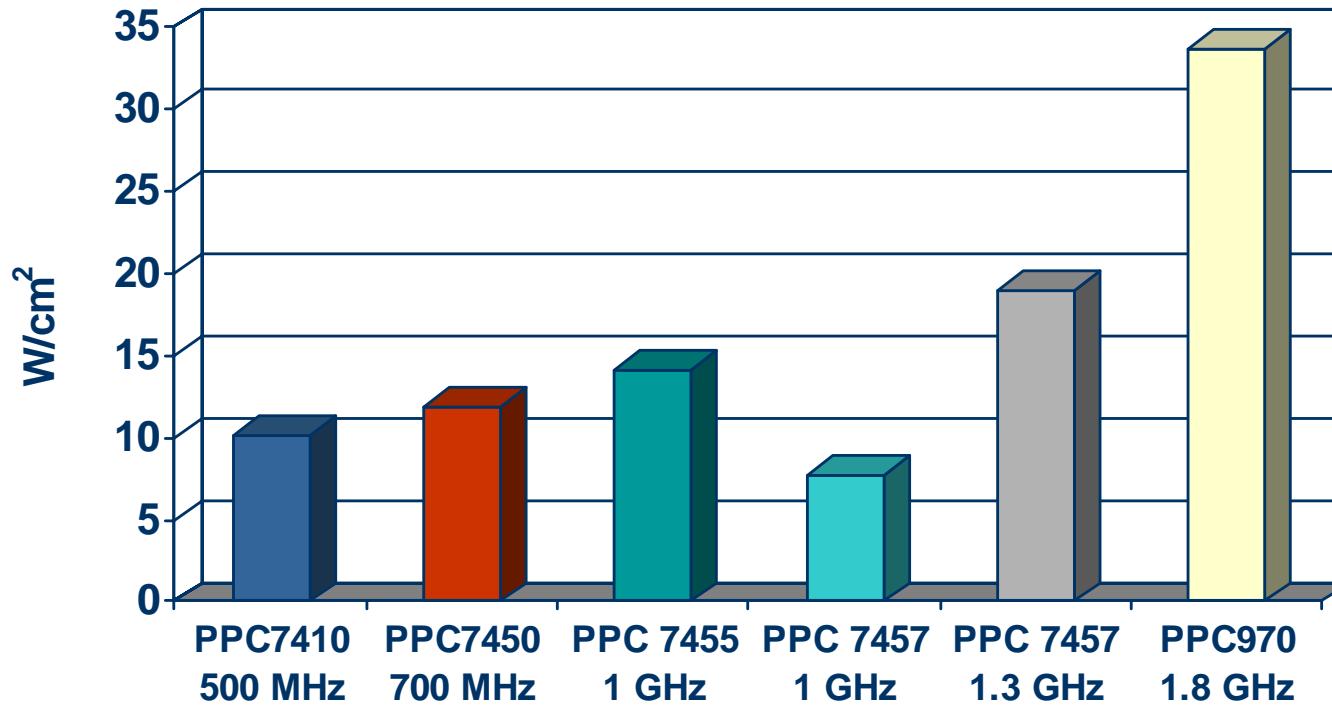
The Ultimate Performance Machine

Objective

- Many high-end deployed military and commercial HPEC applications share a common need to achieve high to very high compute and bandwidth density in the smallest possible volume.
- In addition, deployed military applications layer on additional environmental requirements such as higher levels of shock, vibration, endurance vibration, temperature, and condensing humidity. Each of those adds constraints on the solution space for maximizing compute and communication density.
- Not all HPEC applications can use the same solution due to varying limits on total size or weight and varying levels of ruggedness. For example, the size and weight requirements differ greatly for manned surveillance aircraft, large UAVs, and small UAVs.
- This poster presentation explores two of the many key dimensions to achieving high functional-density systems:
 - ▷ Thermal management
 - ▷ Board real-estate utilization

Processor Power Flux Trends

The power flux (W/cm^2) of the PowerPC processor continues to increase.



Data taken from publicly available sources such as data sheets, physical measurements, and articles

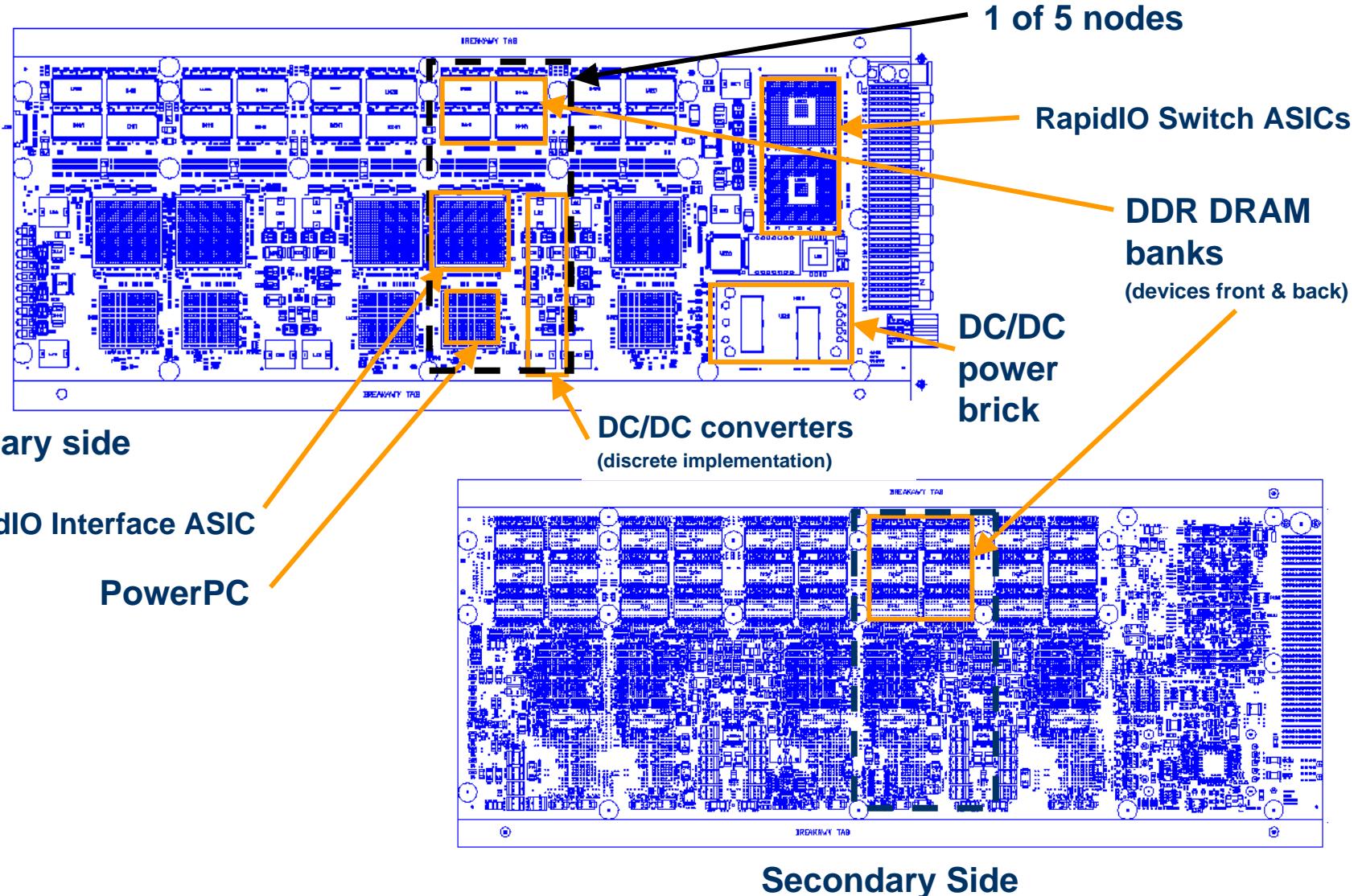
Similar increases are being seen in FPGA compute elements.

1) Thermal Management Study

- Non-processor power is increasing, too
 - ▶ In addition to the increasing power flux of processors, non-processor components in these COTS designs must be thermally managed like never before.
 - ▶ While the PowerPC has been ~80% of a processing nodes total power, it is now closer to 50-60% with the balance of the power dissipation coming from other components of the node; memory, control and interconnect ASICs, FPGAs, and DC/DC power converters.
- Modeling the future
 - ▶ The next few slides contrast the past with the present and future in the areas of thermal management. These CFD modeling results are from a customer-driven study with results correlated against actual similar hardware under operational test.
 - ▶ For visual simplicity, these are modeled as straight-thru air-flow. In a rack-mount system with front-to-back flow where air is taking a 90° turn in and a 90° turn out, resulting temperature may be 10-20+C higher without proper chassis level management.

Processor Board Under Study

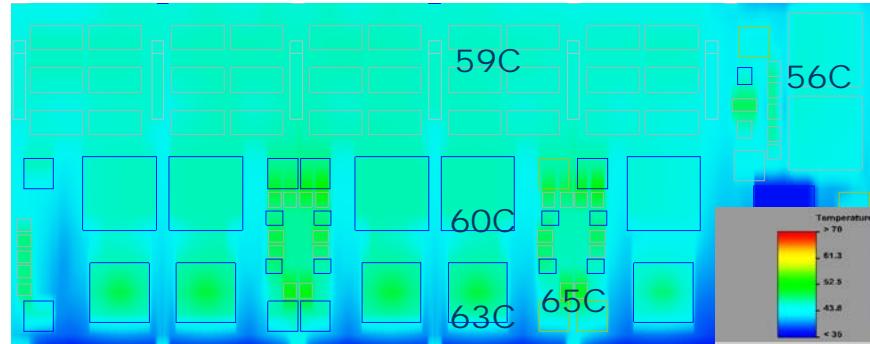
Very high routing and component density



Past Techniques on Past Processors

Past thermal management solutions were fine for commercial designs and had headroom for extending designs to MIL deployed

Models for 8W PowerPC



Commercial: 35C inlet; 5k feet
with "coarsely" managed air



Military: 55C inlet; 10k feet
with "coarsely" managed air

The past:

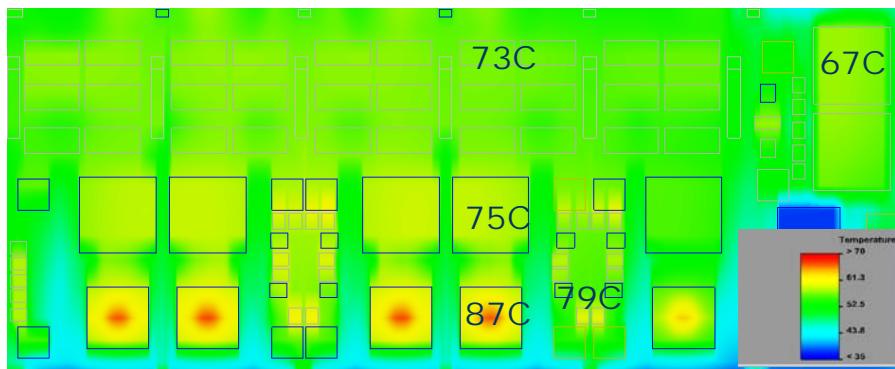
- “Life was good” for commercial environments...
- And, headroom was available for moving to MIL-deployed environments such as 55C concurrent with 10K altitude

Ideal flow shown in this study – temperatures may increase an additional 10-20+C in a non-optimized chassis.

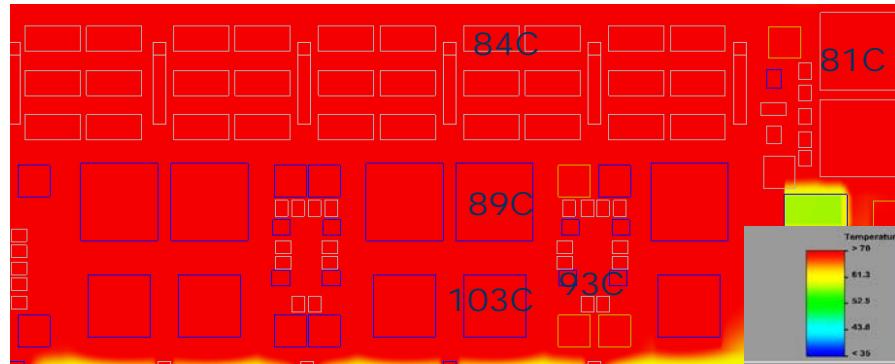
Past Techniques on New Processors

New processors push even the commercial designs to meet spec and make MIL-deployed derivatives unattainable using past techniques

Models for 20W PowerPC



Commercial: 35C inlet; 5k feet
with "coarsely" managed air



Military: 55C inlet; 10k feet
with "coarsely" managed air

Today's problem:

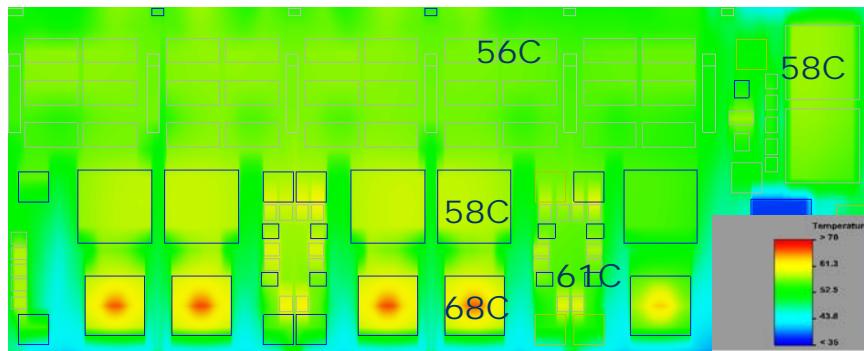
- Even with each processor getting directed air at the inlet temperature (i.e., no preheating), the new generation of processors have little to no margin against their T_j or T_c max temperatures in MIL deployed environments
- New solutions are required!

Ideal flow shown in this study – temperatures may increase an additional 10-20+C in a non-optimized chassis.

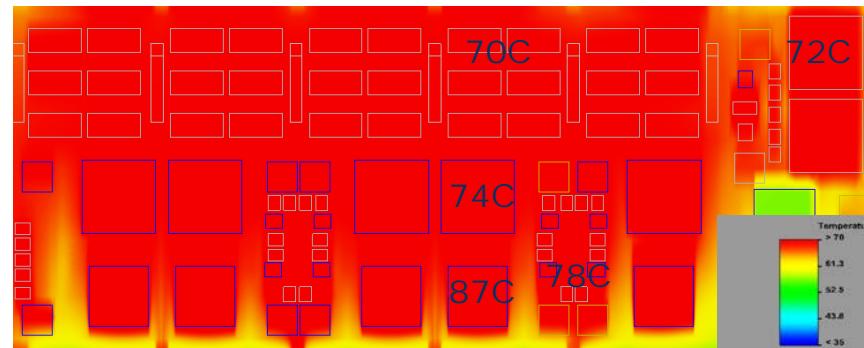
Solution: Finely Managed Airflow

"Finely Managed" air at the board level explicitly shapes and tunes airflow with precise control over impedance drop and impedance distribution over the boards surface

Models for 20W PowerPC



Commercial: 35C inlet; 5k feet
with "finely" managed air



Military: 55C inlet; 10k feet
with "finely" managed air

New solution:

- To allow use of the new generation of processor, memory, ASIC components - new airflow shaping and management techniques are required to achieve similar temperatures to the past commercial designs
- These same techniques enable use of these next-generation components in MIL-deployed environments with only a little headroom

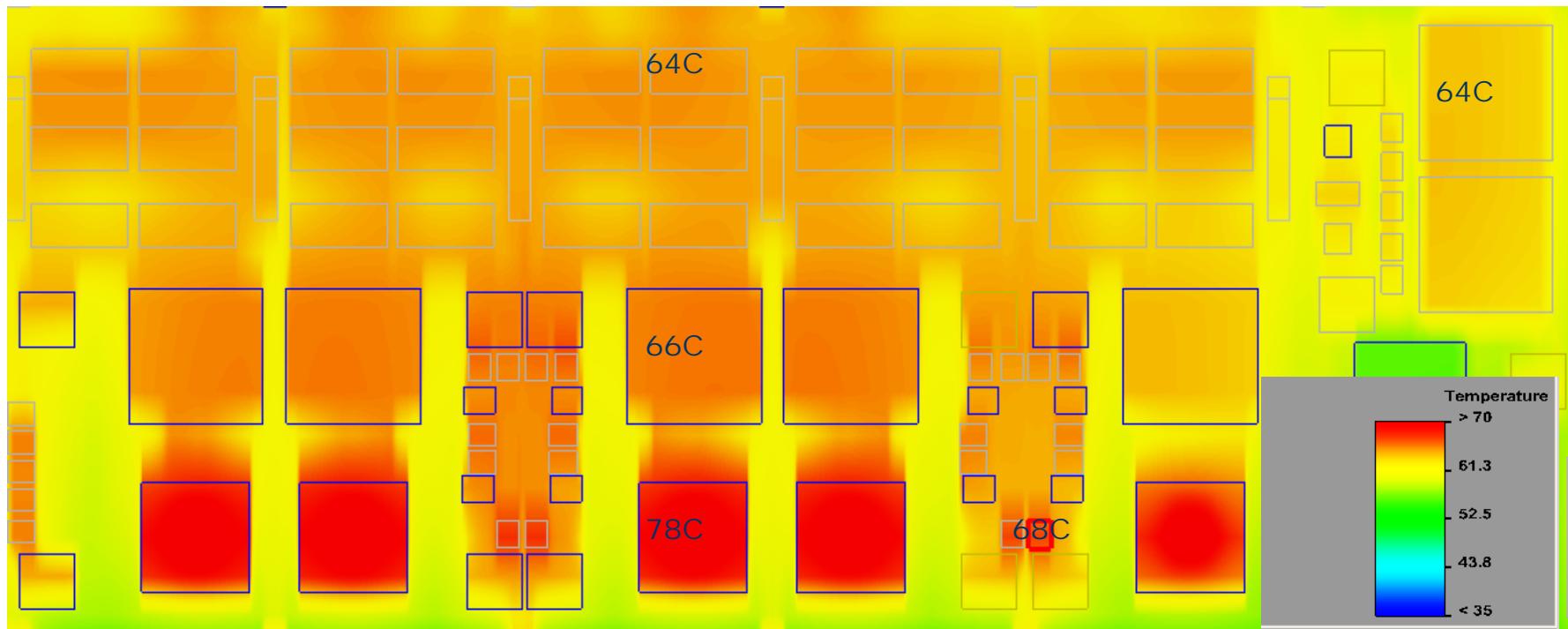
Finely Managed Air Flow at the Chassis Level

Future:

- More aggressive use of these techniques throughout the chassis can result in even more benefit
- *If these improvements weren't achieved, processing and memory frequencies would be reduced by 1/2 to 1/3 – i.e., 400 to 600 GFLOPS (peak) at the chassis level*

Model inputs:

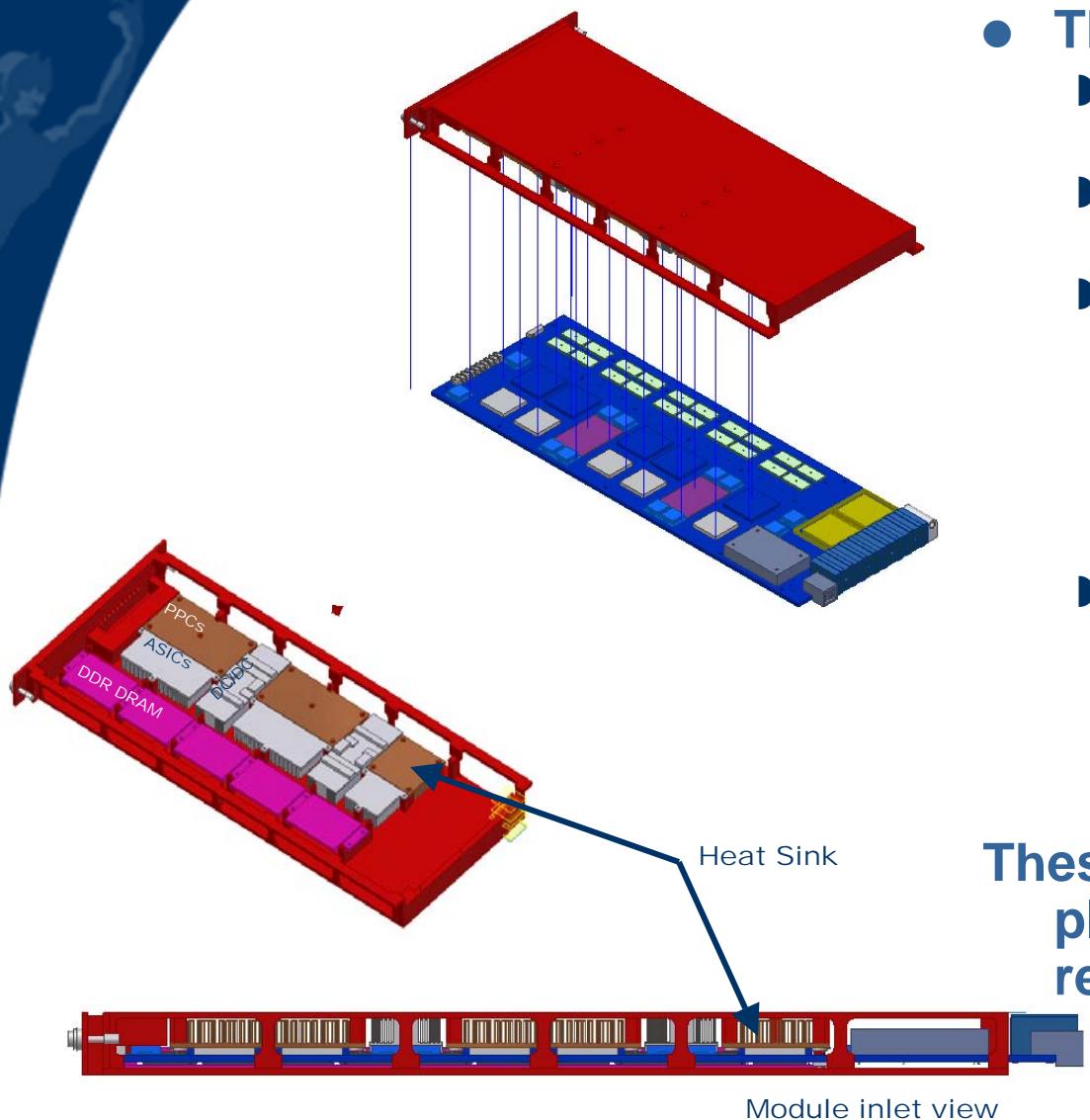
- Processor: 20W PowerPC
- Other devices w/ representative power
- MIL Deployed: 55C inlet, 10K alt.
 - Finely managed airflow, more aggressive throughout the chassis



2) Board Routing and Component Density

- Aside from thermal management, use of board real estate has also become a huge challenge to functional-density.
- The mechanical features such as board stiffeners, rugged heat-sink mounting, and such for MIL-rugged, air-cooled boards takes precious inches² of board space. MIL requirements for board- and system-level endurance vibration and shock pulses drive many of the rugged structural requirements.
- The holes, pads, and keep-outs for these items was now causing significant loss of density – for both internal routing and components.
- Example: Due to routing and placement restrictions, using conventional techniques yielded one less processor node on the board format under study, or 24 less nodes at the chassis level.
Thus approximately 250 GFLOPS (peak) would be left on the table using conventional techniques.
- Innovation for forced air-cooled boards was required and resulted in putting thermal management and structural rigidity features in the Z-dimension via a cover mechanism ...

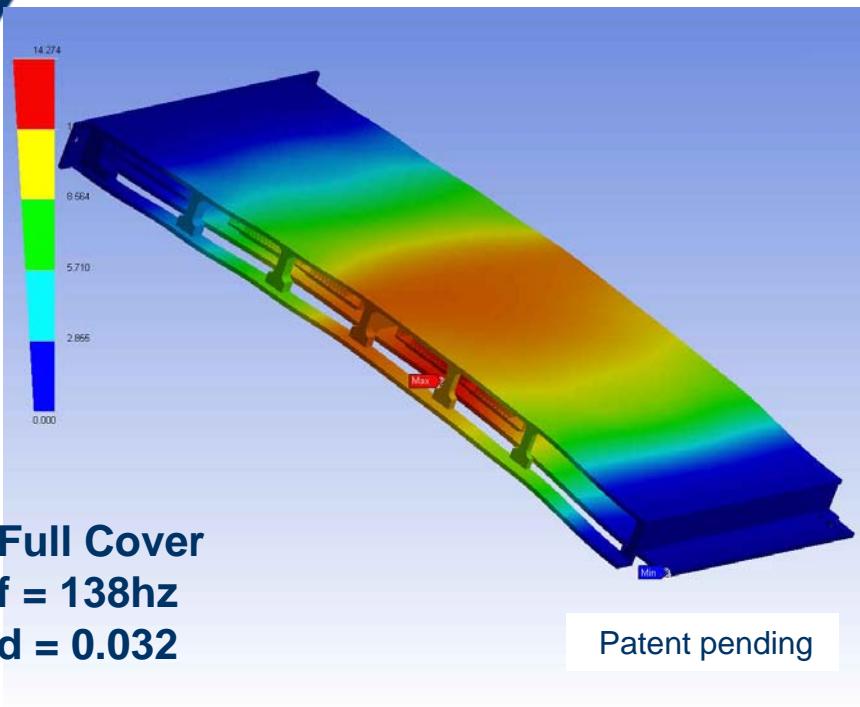
Multi-Purpose Cover Increases Precious Board Space



- The cover provides:
 - ▶ Rugged mounting points for heat sinks
 - ▶ Rugged structural members that the board mounts into
 - ▶ Facilities for finely managed air features
 - In the cover surface itself
 - On the inlet
 - On the outlet
 - At the interface to the card-cage
 - ▶ Finely controlled air plenums are created at a slot-by-slot level at the board and card-cage structures

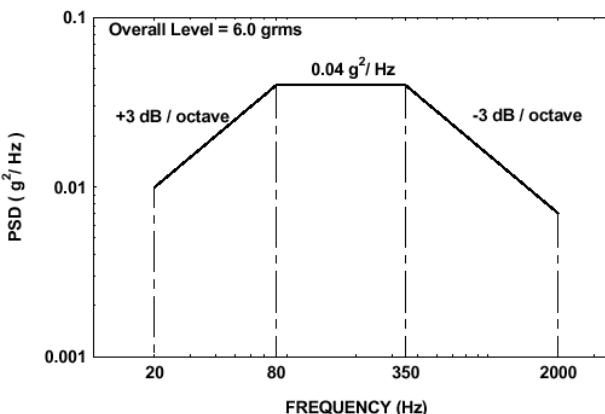
These features can be tuned as platform and application requirements drive needs

Cover Structure Has Improved Performance



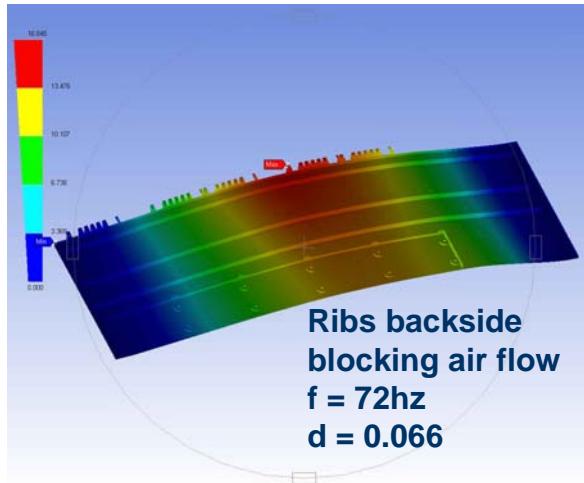
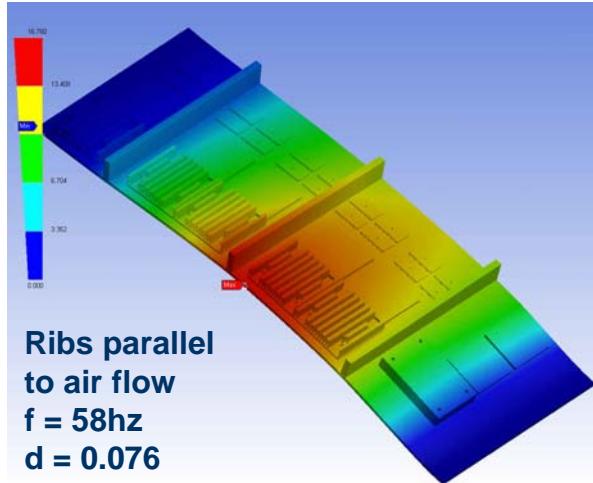
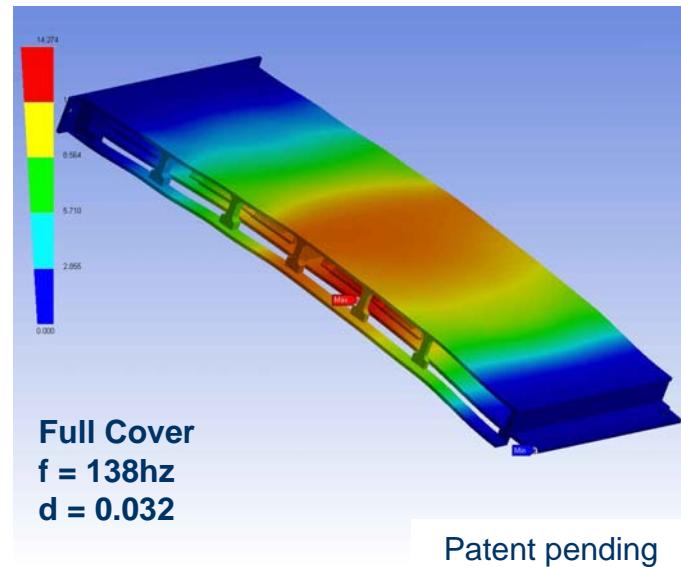
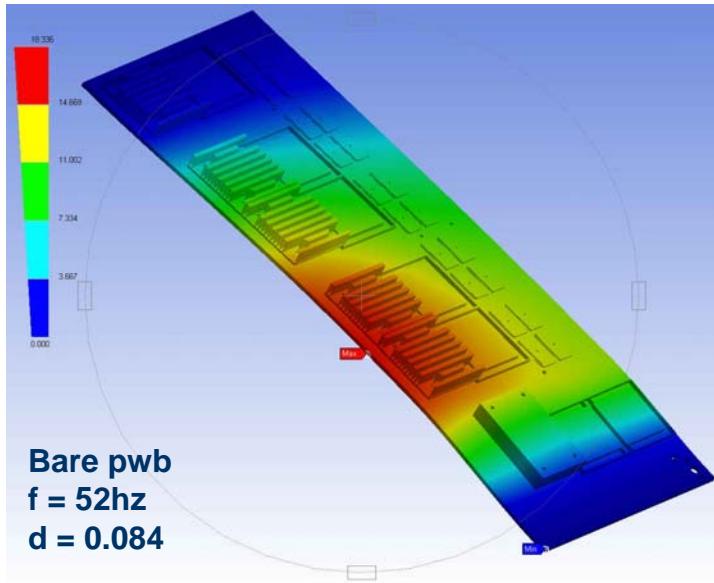
Full Cover
 $f = 138\text{hz}$
 $d = 0.032$

Evaluation done using
NAVMAT P-9492 "Willoughby"
—
A typical random vibration
test for components and PWBs



Performance of full cover using the Z-dimension is similar to or better than other structural methods that take up significant board space and are often at odds with the direction and amount of airflow required.

Comparison Among Other Alternatives For This Module Format



Typical goal for this style of module is:

- $125\text{Hz} < f < 150\text{Hz}$
- $0.025 < d < 0.040$

A Glimpse at a System Solution

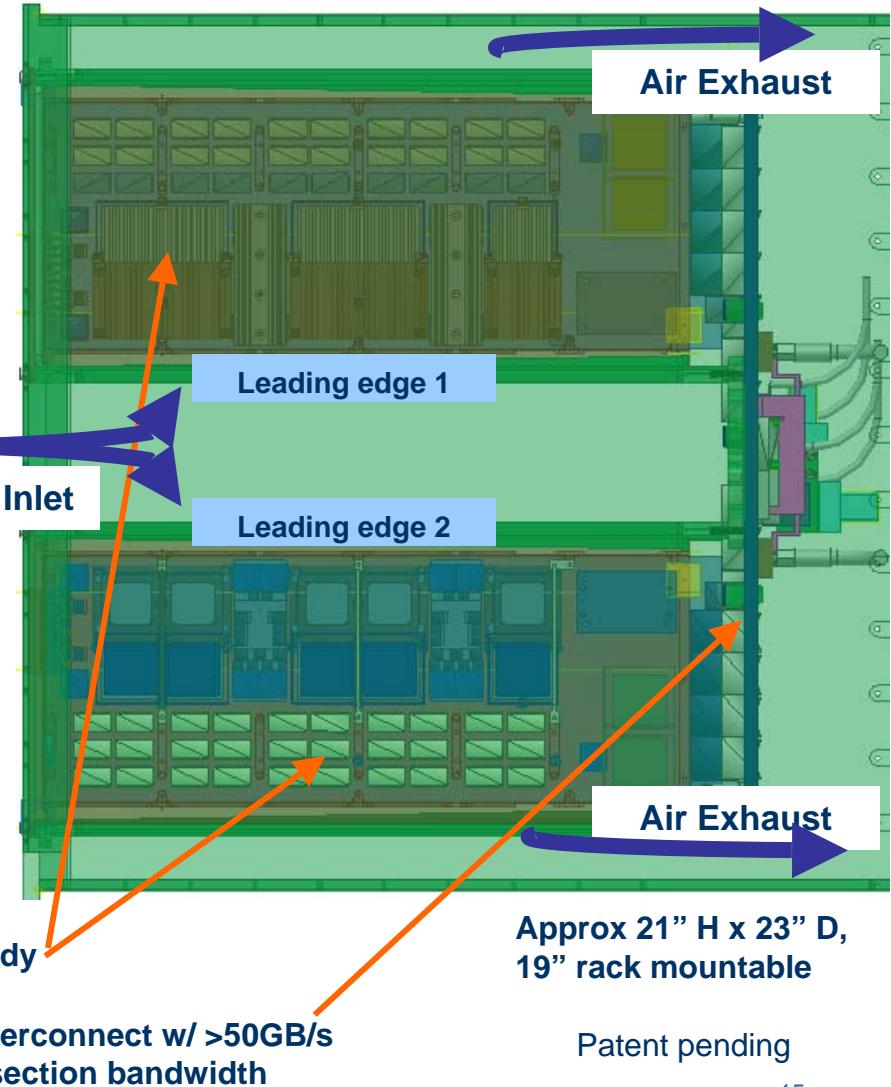
- For a total system solution, solving these present day board-level thermal and structural management isn't the whole issue, chassis-level functional density must complement these solutions.
- The *balanced HPEC TFLOP* under study, in a small volume chassis, would need all of this supporting cast in addition to the processors:
 - ▶ >100 GB total memory
 - ▶ >50GB/s aggregate and bisection backplane interconnect
 - ▶ 10-20GB/s aggregate and bisection inter-chassis interconnect
 - ▶ 10-20GB/s concurrent external I/O
(e.g., streaming sensor data over fiber, VITA-17.1)
 - ▶ 2-3 dozen open standard I/O slots or sites
(e.g., IEEE 1386.1 PMC, VITA-42 XMC)

A Chassis-Level View

- This picture represents the results of one study which determined a means to meet these high functional-density requirements (e.g., “a balanced TFLOP”) in a MIL-deployable HPEC system.
- This arrangement achieves two leading edges of inlet air to enable high-density without having to have “columns” of the hot components, such as processors, heating each other.
- From the previous thermal study presented, it is obvious that the heating effects of these new generation components won’t allow column organizations used in the past.

Processor module under study

Two processor modules mounted vertically sharing the same air intake to minimize vertical space.



Investment in Innovation

- Mercury has invested in innovations toward solving this new class of problems in the COTS MIL-deployed, high-density HPEC application space.
- To date, 8 patents related to the IP for these methods described have been filed. Five of these filings have "notice of allowance" which is the last step to patent registration. The remaining 3 filings are in the last stages of the PTO process.
- The IP represented by these patents will appear in Mercury's future COTS MIL-deployable forced air-cooled HPEC products under development.